Conventions and institutions in coordination problems

Lecq, Fieke van der

Conventions and Institutions in Coordination Problems

Fieke van der Lecq[1][2]

July 1995
SOM theme C: Coordination and Stabilization of Economic Systems

Abstract

This survey article starts with a game-theoretic interpretation of coordination problems that occur in an economy. Three types of coordination games are discussed, in which the degree of coordination versus conflict varies. It is shown that game-theoretic techniques for equilibrium selection or securing the highest pay-off outcome do not always suffice, which raises the need for exogenous information. Coordination norms, such as conventions and institutions, may provide this information. The emergence and persistence of coordination norms as well as the relationship between the type of coordination game and the type of coordination norm are discussed. After a discussion on conventions and rationality, some notions from New Institutional Economics are introduced, in which information and transaction costs are taken as explanations for the existence of institutions. Some applications are given in the last section.

---

[1] Please send correspondence to: University of Groningen, Department of Economics, P.O. Box 800, NL-9700 AV Groningen, The Netherlands, phone +31 50 (3)637182*, fax +31 50 (3)637337*, Email: s.g.van.der.lecq@eco.rug.nl
*: include (3) after October 10, 1995

[2] The author expresses gratitude to Simon Kuipers, Wilfred Dolfsma, Hans van Ees, Harry Garretsen, Maarten Janssen, Remco Kouwenhoven, Bart Nooteboom and Bert Schoonbeek for their comments on earlier versions of this paper without implicating them in remaining errors.

1. Introduction

In an economy coordination problems arise in cases of multiple equilibria. Agents who, because of incomplete information, have to take care of coordinating their actions themselves, can but act on their expectations of other agents' actions (Kuipers et al, 1994). As these other agents act on their expectations as well, decisions become interdependent. In this way, more equilibria can result, depending on the level of mutual expectations. Two examples may serve as an illustration.

Nominal price rigidities can result from a coordination problem. In the New Keynesian theories of nominal price rigidities, explanations are sought in the existence of adjustment costs, which lead to time dependent or state dependent pricing behaviour. Agents have to choose between the costs of adjusting their prices, called the menu costs, and the costs of not doing so. These costs depend on whether or not most other agents adjust their prices, so that agents' decisions are interdependent. This situation brings about two equilibria, namely equilibria in which all agents adjust and equilibria in which no one does. Situations of multiple equilibria leave the question unanswered as to how a particular equilibrium is selected, i.e. how do the agents coordinate their actions in order to reach a particular outcome (Cooper and John, 1988).

Another example of a coordination problem is found in producers' making of investment and production plans, which are partly based on expectations of other producers' plans as well on expected consumers' behaviour. Here, equilibria can exist at higher levels of investment and production, as well as at lower levels of activity. As both situations are equilibria, incentives to move to a higher level equilibrium are lacking. A coordination failure is said to be present if

> "mutual gains from an all-round change in strategies may not be realized because no individual player has an incentive to deviate from the initial equilibrium" (Cooper and John, 1988, p.442).

This may bring about avoidable unemployment and underutilization of production capacity. From a welfare point of view, it is worthwile to study

situations of multiple equilibria that lead to coordination problems.

Multiple equilibria are typical in game theory, where either the rules of the game do not provide for a unique equilibrium outcome or the equilibrium outcome of the game yields an inferior level of utility (Ullmann-Margalit, 1977, p.15). Therefore game theory can be used for the purpose of illustrating coordination problems. This paper gives a game-theoretic description of coordination problems and describes to what extent conventions and institutions as rules of behaviour can solve these problems. In section two coordination problems are presented in a game-theoretic format. Three types of coordination games are discussed, those being pure coordination games, prisoner's dilemma games and bargaining games. It is illustrated that game theory needs expansion beyond the mere construction of equilibria in order to contribute to an explanation of the societal selection of specific equilibria, such as, for example, the production level of an economy. Although the paper is not intended to give such extension to game theory, an informal description of an equilibrium selection mechanism is given in section three. Theories of conventions and institutions are consistent with the game-theoretic framework and provide for a selection mechanism as well. In the theory presented, conventions and institutions are seen as solutions to coordination problems, or outcomes of coordination games, that are compatible with rational behaviour. Conventions and institutions can roughly be described[3] as rules of behaviour in situations that call for coordination. Conventions are often said to be self-enforcing, that is, once they are established, rational individuals will conform to them.[4] Institutions contribute to collective equilibrium selection by means of a sanction mechanism. The persistence of conventions and institutions can be explained from instrumental rationality, because both are means to achieve coordinated outcomes, which in turn contribute to welfare. The emergence of them, however, may as well be historically determined and therefore path-dependent. The process by which conventions and institutions come into existence is elaborated

[3] More formal definitions of both terms follow in section 2.3.

[4] The theoretical debate on this topic will be discussed in section 2.5.

in section three. In section four the relationship between conventions and information is explored within the context of Neo Institutional Economics. Conventions may be seen as information carriers, since they arise in situations of imperfect information. Section five gives examples of conventions and institutions as possible solutions to coordination problems in an economy. The sixth section concludes.

2. Coordination games and coordination norms

"A game is a situation in which a number of individuals or players interact, and in which the outcome for each of them depends not only on what he or she chooses to do, but also on what others choose to do" (Sugden, 1986, p.9). The game format provides a useful tool for analysing human interactions, because it explicitly describes the alternative choices of individual agents as well as the payoffs that result from the combination of choices made by the players of the game. As some combinations of choices result in higher pay-off outcomes than do other combinations, games provide an adequate model of coordination problems.

In game theory, instrumental rationality is assumed, being the selection of actions by agents according to the criterion of their contribution to the realization of the agents' maximal utility. Instrumental rationality implies an optimal use of means to realize ends (Hargreaves Heap, 1989, p.4). An instrumentally rational agent will always use possibilities to increase his profit or utility.

2.1 Game Theory and Coordination Games

In game theory, two often used criteria for equilibrium selection are dominance criteria and the concept of Nash equilibrium. In a Nash equilibrium each player's strategy is at least as good a reply to the strategies of the other player(s)[5] as are all alternative strategies available to him. For a repeated game this implies that

[5] For reasons of simplicity, the remaining text is restricted to two-person games.

players, once they are in a Nash equilibrium, do not want to change their strategy. It is rational for them to stick to it, because this maximizes pay-off to them given the strategy of the other players. Pure coordination problems are characterized by multiple equilibria that result in equal pay-offs and non-coordination equilibria with lower pay-offs to the players. In pure coordination problems a convention may serve as an indicator of what equilibrium to select so that all players choose the same equilibrium strategy and end up in a coordination equilibrium. Therefore, once a convention has been established, players tend to conform to it (Lewis, 1969), which may become the obvious thing to do.[6] In that case ".. if there is an obvious way to play the game, then it is necessary that this is a Nash equilibrium" (Kreps, 1990, p.410). From this point of view, a convention has the same characteristics as a Nash equilibrium in terms of agents conforming to it, given that others do so as well.[7] To sum up: in case the pay-off criterion fails to discriminate between equilibria, the obviousness of a particular solution serves as a selection criterion that enables the players to coordinate.

Pareto-dominance and risk-dominance are useful criteria for ranking the equilibria and eliminating some of them, until one or a few remain. A combination of both dominance criteria is given by Harsanyi and Selten (1988), who present a method of equilibrium selection that is claimed to yield a unique outcome in all games. Their method can be considered an extension of the concept of Nash equilibrium (Van Damme and Heertje, 1994, p.940). However, in situations of multiple Nash equilibria, problems arise. As Sugden (1989) points out, if rationality leads to a unique outcome, then conforming to a convention cannot be rational. With Sugden, I tend to believe the opposite is true: "if it can be rational to follow a convention, the claim that every game has a unique rational solution must be false" (Sugden, 1989, p.86, see also Furth, 1989 and section 2.5). An extension of the Nash equilibrium concept may

[6] Section 2.5 deals more extensively with this issue in the context of rationality.

[7] In section 2.3 Lewis' formal definition of a convention is given. The third part of it is especially relevant in this context.

therefore be insufficient to guarantee the selection of a unique outcome. As the selection theory of Harsanyi and Selten rules out all but one of the (strict) Nash equilibria, it cannot rely on rationality alone and has to be based on a theory of behaviour as well, which is external to the game (Friedman, 1994, p.85). Therefore, it can be inferred that games do not always generate one unique equilibrium, nor the outcome with highest pay-offs to the players. This is illustrated for three types of coordination problems.

Coordination games can be ordered on a scale with games of pure conflict[8] at the one extreme and games of pure coordination at the other extreme (Schelling, 1960, p.84). Most games are neither games of pure conflict nor games of pure coordination, but are a mixture of opposition and coincidence of interests (Lewis, 1969, p.14). The three types of games introduced below can be put on this scale as follows. First, pure coordination games are identical to the one extreme of the scale. As Lewis defines them: "Games of pure coordination, in which the agents' interests coincide perfectly, are games in which the agents' pay-offs (perhaps after suitable re-scaling) are equal in every square" (1969, p.14). Second, prisoner's dilemmas are close to, but not identical with, the other extreme, because there is an element of coordination in it, as one combination of strategies yields higher payoffs to both players than the others do. By coordinating their actions, the players could reach that higher pay-off and so move from a low pay-off outcome to a higher one. The degree of conflict dominates the element of coordination, because the non-coordination outcomes have the highest pay-off to one of the players. Third, bargaining games can be

[8] Pure conflict games are also called zero-sum games, because the loss of one player is the gain of the other(s). Strictly speaking, as in these games no coordination element is present, the extreme itself falls outside the group of coordination games. An example of a zero-sum game is given by Ullmann-Margalit (1977, p.79):

	player II	
player I	1 , -1	0 , 0
	0 , 0	1 , -1

In this pay-off matrix and in the following ones the pay-offs to player I are given first.

put somewhere in the middle of the scale, because the conflict element and the coordination element balance each other. In bargaining games the equilibria cannot be Pareto ranked, because moving from the one equilibrium to the other has a redistributive effect. However, coordination may improve total pay-off and therefore be worthwhile. As, to some extent, in all three types of games elements of the coordination problem are present, they all can be called coordination games.

First, in a pure coordination problem players have to find a way of choosing the same (pure) strategy in order to reach one of the Pareto superior equilibria. An example of a pure coordination game is:

(table 1)

		player II
player I	1 , 1	0 , 0
	0 , 0	1 , 1

in which player I can choose between playing 'top' or 'bottom' and player II can play either 'left' or 'right'. If pre-play communication is allowed, players can agree on a strategy. Otherwise, they must conjecture the strategy their opponent will play. In case the game is played repeatedly, players can base their expectations on history and a convention can come into existence. This will be discussed in section 2.4.

Second, in a Prisoner's dilemma game, the Nash equilibrium is Pareto inferior, as the next example (Kreps, 1990, p.504) shows:

(table 2)

		player II
player I	5 , 5	-3 , 8
	8 , -3	0 , 0

The players are inclined to play 'bottom, right' and so end up with payoffs '0,0'.

A binding agreement can enforce the strategy combination 'top,left' so that the Pareto-superior outcome of '5,5' is realized. If the Prisoner's Dilemma game is played repeatedly, a reputation of playing a tit-for-tat strategy may serve as an instrument for attaining the Pareto-optimal outcome, even if a binding agreement is lacking (see for instance Furth (1993), Kreps et al. (1982) and Rasmusen (1989)). Although reputation building is a phenomenon that can be captured within the game context, it is of little use to us, as reputations can only be built if the game is repeatedly played by the same players. If the players who oppose each other are members of a larger community and seldom meet again, reputation building is fairly difficult. This makes many forms of trade anonymous (cf. Sugden, 1986, p.35).

Third, a bargaining game,[9] which is a game of conflict with multiple equilibria (Hargreaves Heap, 1994, p.40). For instance, it may look like:

(table 3)

		player II
player I	6 , 3	2 , 2
	2 , 2	4 , 4

This game differs from the PD-game by the impossibility of Pareto-ranking the equilibria. Here, 'top,left' and 'bottom,right' are both Nash equilibria without one being Pareto-superior to the other, but total pay-offs being largest in 'top,left'. If a binding agreement can be reached, for instance an agreement by which player I pays one-and-a-half to player II in order to have him choose 'left', the outcome with the highest total pay-offs can be attained. If such agreement cannot be reached, the players may choose 'top,right' and end up with '2,2'. If the game is played repeatedly, it may be the case that players used to choose 'bottom,right' and cannot get to the 'top,left' equilibrium. In that case a norm external to the game may help. In an economy, bargaining games are

[9] Bargaining games are also called negotiation games, chicken games and mixed-motive games.

important in situations of rent-seeking behaviour, because this behaviour may reduce total output. Solving the bargaining game can contribute to a higher level of total output by discourageing rent-seeking behaviour (Hargreaves Heap, 1994, p.40).

Besides the distinction between coordination and conflict, games can also be divided in cooperative games versus non-cooperative games. "If both communication and binding contracts are possible, the game is called a cooperative game; if no communication is possible, the game is called non-cooperative" (Schotter, 1981, p.15). The latter part of the definition stems from the necessity of communication for making a binding contract and so playing a cooperative game.

The distinction between cooperative versus non-cooperative games can be combined with the coordination-conflict scale introduced above. Wubben (1993, pp.217-223) puts both scales orthogonal to each other, so that a 2x2 matrix can be made. We will discuss the four cells of this matrix two-by-two. First, in a cooperative game, coordination is redundant. If a binding agreement forces players to choose a particular equilibrium strategy, they need no longer coordinate on it. It is interesting to note that the findings of Schelling and Lewis underpin the obsolescence of a coordination-cooperation combination of game types, because they have shown that an explicit agreement is not a necessary condition for successful coordination (by way of a convention) to exist (Ullmann-Margalit, 1977, p.76). Second, in a game with neither coordination nor cooperation, a binding agreement cannot provide for the attainment of the outcome with the highest collective pay-off. The conflict results in a Pareto-inferior outcome. Taking these two cases together, it can be concluded that in a cooperative game the presence or absence of coordination does not matter, whereas in a game of conflict, the possibility of a binding agreement is essential. This line of reasoning enables me to reduce the orthogonal combination of both scales to a one-dimensional combined scale, with coordination games at the one end and cooperative games at the other end. In my view, the distinction between cooperative and non-cooperative games runs parallel, though inversely, to the

coordination-conflict scale:

coordination	*conflict*
non-cooperation	*cooperation*

(figure 1)

In this figure, the remaining two combinations are shown. In the third one, situations that call for coordination need not demand cooperation. An example is found in the use of a medium of exchange, because the pay-off in terms of transaction efficiency is higher for all agents if they use the same medium of exchange, but they tend to do so without the necessity of a binding contract. Later I will show that conventions, being solutions to coordination problems, tend to be self-enforcing, so that a binding agreement is redundant. In the fourth case, cooperation may provide for a binding agreement in a problem of conflict. However, rational agents will only cooperate if they profit from the existence of a binding agreement or are forced by threat of a sanction to enter one, so that at least a small element of coordination remains, just like in PD-games. In an economy, such situations occur in negotiations, cartels and co-makership relations. Here, the binding agreement acts as an institution: it solves the coordination problem and enforces compliance by the sanctions it contains. As the original four cases are reduced to two, a reduction of terms is also possible. In the remaining text, the scale of figure 1 is reduced to coordination games at the one side and cooperative games at the other side.

2.2 Coordination norms

In the above three types of coordination games the resulting pay-off for individual players depends on the action of the player himself as well as on the action of the other player. Without further additions to the game setting, a prisoner's dilemma game results in a low total pay-off outcome, a bargaining game may do so, and a coordination game does not lead to a determinate outcome. These situations can be seen as problems, which cannot be solved

within the game. Therefore something external to the game needs to be supplied, e.g. a norm.[10] A norm is defined as follows: "A social norm is a prescribed guide for conduct or action which is generally complied with by the members of a society" (Ullmann-Margalit, 1977, p.13). According to the type of problem, the norm takes a different form: a convention, an institution or a combined form. All three types of norms are common in that they change the pay-off matrix in such fashion that one strategy combination becomes more attractive than others. They will now be treated in sequence.

In a pure coordination game, the aim of selecting an equilibrium is shared by the players and creates interdependency between them. If there is indifference with regard to the pay-offs, so that players do not mind which equilibrium they end up in as long as they end up in one, then players have the common interest of picking out the same equilibrium (Ullmann-Margalit, 1977, p.115). If the equilibrium solutions clash, then both players also have an interest in choosing the same equilibrium strategy.[11] In both cases, the strategies of the players depend on each other. Strategic interaction, being the interdependence of strategies may lead to infinite regress and cause players to make no decision at all.[12] In order to end this stalemate 'something' outside the game should give the players a clue as to what their fellow players will choose to do.[13] This

---

[10] Of course one can also supply a dictator, but then the rationale behind doing so is left unexplained. The attraction of a norm is that its existence can be explained without changing the number of decision making agents, so that it is compatible with methodological individualism (see also section 3.1).

[11] A classic example is the coordination of driving lanes: if one player chooses the strategy of driving at the left-hand side and the other player chooses the strategy of driving at the right-hand side, then both solutions clash (as well as may both players).

[12] 'No decision' can also mean 'no decision to change' so that inertia results. This may occur if producers have to decide whether or not to issue new prices or to change wages. Another example is given by investors who, in situations of uncertainty, may decide not to change their investment behaviour (Keynes, 1936, pp.152-3).

[13] The coordination of actions does not always mean choosing the same action, as can be illustrated with the coordination of traffic at crossroads, whereby one party gives way to the other. In other words, not all coordination games are unanimity games, in

’something’ can be a pure coordination norm: a convention. Conventions help players choose one out of some coordination equilibria.

Prisoner’s dilemma type of norms make players choose the strategy combination with the highest total pay-off. Because this is not a Nash equilibrium, players tend to deviate and will only comply with it if it is accompanied by sanctions. The sanctions can either function as a restriction on the number of strategies open to the players, for instance by prohibiting the choice of the degenerate strategy combination, or by changing the pay-off structure in such a fashion that the players choose the strategy combination with the highest total pay-off[14] (Ullmann-Margalit, 1977, p.37). Examples of such norms can be found in morality, culture and religion.

Bargaining game norms are a relatively difficult case, because they have a hybrid character. If the players choose the one with maximal total pay-off, no norm is needed[15]. However, if they select the outcome with less total pay-off, for instance the Nash equilibrium ’bottom, right’ in table 3 above, a norm is necessary to have them choose the strategy necessary for attaining the equilibrium ’top, left’. A sanction may be necessary to make them sure the other player will change his strategy as well. Further, as is also illustrated above, the move to the equilibrium with maximal total pay-off may be accompanied by redistributive effects. In that case sanctions are necessary because incentives diverge (Hargreaves Heap, 1994, p.41). Another possibility is negotiations that secure compensation of the losing party by the winning one, as was illustrated above.[16]

---

which there is positive pay-off only when all players play the same strategy (Furth, 1989, p.737).

[14] A nice example is found in the IMF board, where consensus is an enforceable norm, with voting as the sanction mechanism (Witteveen, 1994).

[15] For simplicity, the attainment of maximal total pay-off is considered more desirable than are preferences with regard to distribution.

[16] Here again, legal intervention or (benevolent) dictatorship are possibilities that can be thought of, but do not explain anything.

To sum up, in coordination games, multiple equilibria call for a coordination of strategies, and in cooperative games, total pay-off increases if a binding contract resolves the inherent conflict between players. A line drawn parallel to the one above indicates the accompanying solutions to the respective coordination games:

coordination	← game type →	*cooperation*
conventions	← solution type →	*institutions*

(figure 2)

As in an economy many confrontations are ad hoc and anonymous, an explicit agreement is seldom made, which increases the relevance of studying games with a high degree of coordination, which are posited at the left-hand side of the scale.

The three types of coordination norms differ in two aspects. First, a prisoner's dilemma norm and a bargaining game norm are not arbitrary, as only one outcome has the highest total pay-off, whereas a pure coordination norm is. Second, a prisoner's dilemma norm is socially first-best, but individually second-best, while a coordination norm is socially as well as individually first-best (Ullmann-Margalit, 1977, pp.118-119) in terms of pay-offs. The rating of a bargaining norm depends on the player's situation.

2.3 Taxonomy and terminology

Before linking the game-theoretic interpretation of coordination problems to the notions of conventions and institutions it should first be clarified what is meant by both concepts. Many authors use different interpretations of these terms. In this article the framework of terminology will be a combination of the one used by Ullmann-Margalit (1977) and that of Schotter (1981), because it is compatible both with the games presented above and with the (institutional) economics literature cited below. Norms that solve repeated coordination problems can be divided into the class of conventions or the class of decrees: "Two types of such

norms are distinguished: (1) accepted and established solutions to past recurrent coordination problems which - with time - assume the status of norms; (2) solutions to novel recurrent coordination problems which from the outset are being decreed as norms. The first type is then called ‘conventions’, the second ‘decrees’." (Ullmann-Margalit, p.76). The latter type of norms is often called institutions, which makes me do so as well, in order to avoid unnecessary differences in terminology. Others, such as Schotter, define institutions in such a way that conventions are a subgroup of institutions: "..social conventions: regularities in behaviour which are agreed to by all members of a society and which specify behaviour in specific recurrent situations" (Schotter, 1981, p.9) and "A social institution is a regularity in social behaviour that is agreed to by all members of society, specifies behaviour in specific recurrent situations, and is either self-policed or policed by some external authority" (ibid., p.11). In this study, conventions are institutions without external sanctioning mechanism that makes people comply with it, or, the other way around, conventions are self-enforcing whereas institutions need not be so. For a more formal definition of conventions Schotter uses the one given by Lewis (1969, p.58) and creates an analogous one himself for institutions:

> "A social convention. A regularity R in the behavior of members of a population P when they are agents in a recurrent situation S is a convention if and only if it is true that, and is common knowledge in P that, in an instance of S among members of P,
>
> 1. everyone conforms to R;
> 2. everyone expects everyone else to conform to R;
> 3. everyone prefers to conform to R on the condition that the others do, since S is a co-ordination problem and uniformly conformity to R is a co-ordination equilibrium in S." (Lewis, 1969, p.58)
>
> "A social institution. A regularity R in the behavior of members of a population P when they are agents in a recurrent situation Γ is an institution if and only if it is true that and is common knowledge in P that
>
> 1. everyone conforms to R;
> 2. everyone expects everyone else to conform to R; and

3. either everyone prefers to conform to R on the condition that the others do, if Γ is a co-ordination problem, in which case uniform conformity to R is a co-ordination equilibrium; or

4. if anyone ever deviates from R it is known that some or all of the others will also deviate and the pay-offs associated with the recurrent play of Γ using these deviating strategies are worse for all agents than the pay-off associated with R." (Schotter, 1981, p.11)

The strange thing with this definition is that it is said to incorporate a sanctioning authority, because Schotter summarizes his definition as follows:

> "A social institution is a regularity in social behavior that is agreed to by all members of society, specifies behavior in specific recurrent situations, and is either self-policed or policed by some external authority" (Schotter, 1981, p.11).

However, an external authority need not itself deviate from the norm in order to sanction the deviator, as is suggested by (4) above, or even participate in the game. Further, from a methodological point of view, it may be useful to distinguish institutions that are created by the players from institutions that are installed by a body external to the players, such as a government or a benevolent dictator. For now, I will combine both types, because the presence of any kind of sanctioning mechanism is the characteristic that distinguishes institutions form conventions. To summarize, in the remainder of this article, a norm is said to be a convention if it is self-enforcing and it is said to be an institution if an enforcing authority such as a sanctioning mechanism is needed to support it. For example, a cartel is considered an institution while the annual issuing of new price lists in some sectors is a convention.

Unfortunately, some authors use the term of institutionalized convention, for phenomena such as firms and markets. I see two reasons for not doing so. First, by using this term they stress the fact that the convention has become a separate entity. In accordance with Ullmann-Margalit (1977, p.17) I consider this a risky approach, because it suggests that the convention itself may interact with agents (see also section 3.1). In this respect, I prefer the methodological individualistic approach of game theory, in which conventions are a result of

interactions between agents instead of a separate entity. Second, the binding element of a convention norm is conceptually different form the sanction mechanism behind an institution, and therefore the term institutionalized convention is unnecessarily vague. In line with this, another terminological diffusion comes up. Although most authors put firms and markets under the heading of institutions and often do so for good reasons, in the context of this article the motive for confirming to the norm is considered important, which makes me stick to the terminology that was introduced above. To sum up: a convention is a coordination norm that is self-enforcing, whereas an institution is a coordination norm that is accompanied by a sanction mechanism.

2.4 From coordination problems to coordination norms

Now that the terms convention and institution are defined and their relationship with coordination problems is introduced, I take a closer look at the mechanism behind the evolutionary emergence and persistence of coordination norms. Referring to the scale on which coordination problems can be ordered according to the degree of coordination versus conflict, my treatment will be most explicit for pure coordination problems. The more conflict is involved, the less a norm will be self-enforcing. Then a more or less severe sanction mechanism is needed and the norm will be of the institution type. I focus on pure coordination problems and conventions that may result, because I consider the self-enforcing character of a convention theoretically more interesting than the external enforcement mechanism within an institution, which is simply decreed.

If a coordination problem emerges, agents will try to solve it in order to reach a higher pay-off.[17] Then, supposed that the agents somehow succeed in coordinating their strategies, if the problem occurs again, they tend to use the same solution in order to avoid the unnecessary costs of again solving the problem[18]. The repetition of the problem creates a tacit agreement on the choice of a particular solution. Conventions emerge as "accepted and established

[17] We assume utility to be increasing with pay-off.

[18] In section 2.5 this statement is subject to discussion.

solutions to past recurrent problems which - with time - assume the status of norms"[19] (Ullmann-Margalit, 1977, p.76). A crucial point in the mechanism is the transmission of the knowledge of the earlier chosen solution to agents who were not involved in past occurrences of the problem. For these 'new' agents, the convention obtains its normative character if they start expecting that others expect them to behave according to the convention. This expectation formation may start from observing the way others behave, or by way of a convention having become habitual or part of culture. Knowing these expectations, a 'new' agent tends to conform to the convention, because it enables him to enjoy a larger pay-off as compared to not conforming. In case of a convention, the pay-off is larger because the agent finds himself in a coordination equilibrium, whereas in case of an institution an agents who complies to the norm avoids the sanction. So, given the expectations of others, the new agents find it in their own interest to follow the norm. The convention that thus comes into existence is self-enforcing, because in a convention the agent prefers himself to conform if others do so as well above he alone not conforming. In this way the repeated behaviour becomes binding. If a 'newcomer' complies to the norm because others prohibit him acting otherwise, albeit for his good as well, then the norm is an institution.

In game theoretic terms, conforming to a convention can be a so-called evolutionary stable strategy: a strategy that is a best reply to itself and to other strategies[20] (see e.g. Sugden, 1986, pp.25-29). The adjective 'evolutionary' indicates that the most successful strategy survives at the expense of less suitable ones. In a pure coordination game, exactly what is the most successful strategy cannot be determined because more strategies are equally successful. Once one

[19] Recall the definition of norms given in section 2.1 above.

[20] "A strategy in a Nash equilibrium is evolutionary stable when there is another strategy that works as well against it as it does against itself. Then this strategy should do (strictly) better against the other strategy than that other strategy does against itself." (Furth, 1993, p.362) Roughly, an evolutionary stable strategy can be interpreted as a dynamic counterpart of the Nash equilibrium. However, not all Nash equilibria are evolutionary stable (Sugden, 1989, p.91).

of them is chosen, it will be the best one because all players choose it. From then on, repeating the chosen strategy[21] is cheapest in terms of coordination costs, being the information and transaction costs that are incurred by the coordination of strategies. A stalemate emerges: a strategy is successful because it is chosen and it is chosen because it is successful. To break the stalemate, something outside the game is needed to indicate the strategy to be chosen. I will come back to this 'focal point' in section three. Not only strategies are subject to an evolutionary process, but conventions themselves are as well. A convention is successful if it can easily be applied in situations that are analogous to the situation from which it emerged (Sugden, 1986, p.51). In this way the domain of a convention, being the range of coordination problems for which it serves as a solution, can become smaller or larger depending on the relative success of the convention as compared to other conventions.

With regard to conventions, in which the binding aspect of the norm is endogenous, two aspects of the transformation of repeated behaviour into a binding pattern of behaviour can be worked out more explicitly. Firstly, why does a norm become binding? Secondly, how does a norm become binding? Two reasons can be given as to why a norm becomes binding. First, as is described above, a norm enables agents who find themselves in a coordination problem to secure the attainment of a coordination equilibrium. Second, they do so at lowest information cost, for three reasons. First, they inform agents who face the coordination problem for their first time on how to solve it. Second, the norm gives a clear algorithm about how to deal with a specified type of problems, so that future occurrences of alike problems do not lead to ambiguities on how to behave. Third, a norm helps to make the connection between a particular problem and a particular solution more explicit to all members of society. In this respect, a norm can be seen as a public good, because it satisfies the criteria of non-rivalry and non-excludability. It is not subject to rivalry, because it does not diminish by being used by more agents, and it is non-excludable, because agents cannot prohibit each other's conforming to it. The

[21] Perhaps this can be interpreted as adaptive behaviour.

fact that some solution is the generally accepted one, slightly changes the pay-off structure of the problem to the advantage of the chosen solution (Ullmann-Margalit, 1977, pp.85-87). Strictly speaking, in this way, a pure coordination problem ceases to exist once a convention is established, for the requirement of there being at least two equal pay-off equilibria is no longer satisfied, since one of them has become more attractive by its being chosen before. Put differently, a pure coordination game can only be repeated up till the establishment of a convention, although this is a gradual process. *From then on, it has become a different game*. The question as to how a norm becomes binding has a twofold answer. First, communication of the norm increases the probability of everyone conforming to the convention and so attaining a high pay-off outcome (see also Sugden, 1986, p.41). Rational agents who repeatedly chose a particular strategy to solve the coordination problem will therefore contribute to the diffusion of a norm by communicating it to others, as long as the costs of communication are lower than the marginal pay-off that is gained by coordination. Second, motives that are not necessarily linked to pay-off may support the process of a norm becoming binding. Sugden (1989, p.95) considers the human desire for the approval of others as crucial to the transformation of a regular pattern of behaviour in a binding norm. The desire for approval forms the basis of social pressure to conform, which emerges as people expect each other to conform to the norm. Related arguments are reputation and the feeling of belonging to a group, which also underpin a pressure to conform (Boyer and Orléan, 1992, p.166). Further, these mutual expectations tend to get transformed from theoretical expectations to deontic ones. This means that if an individual agent does not behave in conformity with the norm, the other agents do not only regret this because the pay-off they get is less than it could be, but they also resent it because the agent did deviate *per se* (Ullmann-Margalit, 1977, pp.88-89). The resentment of others influences an agent's behaviour because of the desire for approval (Sugden, 1989, p.95). These non-pecuniary motives are supplementary to the above explanation of norms in terms of solutions to coordination problems, but not necessary for the underpinning of their existence. They merely contribute to the change of the pay-off structure once a convention has come

into existence. From this underpinning of the process of the emergence of coordination norms, it turns out that the role of expectations that agents form about other agents' behaviour and their expectations is very important.

Once a (pure) coordination problem is solved and a convention is established, individual agents tend to conform to it, because it makes the agents involved end up in one of the coordination equilibria, which have a higher return than other outcomes.[22] In this fashion conventions tend to persist. Even if an individual agent does not see the point of using exactly that norm, he knows he is expected to conform to it, and, what is more, the fact that he expects others to conform to it makes his conformity to the norm rational behaviour. However, the norm itself needn't be instrumentally rational in terms of pay-off as more strategy combinations may result in the same pay-off. The utility that is gained by agents who conform to a convention has to be differentiated from the welfare a convention may bring to society as a whole.[23] In a pure coordination problem, a convention does not emerge because the behaviour it prescribes is supposed to be welfare improving, but only because it is a solution to a coordination problem and thereby welfare improving. Therefore, the contents of conventions may or may not contribute to social welfare. The fact that a convention exists, however, does contribute to welfare as it enables society to reach a coordination equilibrium (Sugden, 1986, p.166-167). In terms of instrumental rationality, the conventional behaviour is a means to reach the end of maximal utility given the context of the coordination problem. As this goal can only be reached by the simultaneous choice of the same strategy combination, the convention serves as a facilitator for coordinating actions in terms of lowering coordination costs.

The convention then can be seen as a rule that prescribes actions of agents in situations that call for coordination. In this context, Rowe (1989, p.4) distinguishes act-individualism from rule-individualism. In cases of act-

[22] Whether or not this tendency can be explained from rationality, is discussed in section 2.5.

[23] This holds for all three types of coordination games.

individualism the isolated action is the unit of rationality, whereas in cases of rule-individualism the rules of action are rational. "A rule of action is rational if, by following that rule, an agent maximizes his expected utility" (Rowe, 1989, pp.4-5). In this sense a game strategy can be seen as a rule of action. Further, Rowe defines society as agents following rules (ibid., p.6) and stating that "what we call social institutions[24] are in fact nothing more than agents rationally following rules of action, and being believed by other agents to do so" (ibid., p.5). The transformation of a rule into a convention takes two steps:

> "... a rule is likely to acquire moral force if it satisfies two conditions:
>
> 1. Everyone (or almost everyone) in the relevant community follows the rule.
>
> 2. If any individual follows the rule, it is in his interest that his opponents - that is, the people with whom he deals - follow it too.
>
> Any rule that is a convention necessarily satisfies a third condition:
>
> 3. Provided that his opponents follow the rule, it is in each individuals interest to follow it." (Sugden, 1986, p.166).

This description bears close linkage to the definition of a convention given by Lewis and Schotter that is cited above. Step two begs the question insofar as it leaves unexplained why the player's opponents should follow the rule. The answer that easiest comes to mind is that all agents are assumed to have the same interest, so that step 2 holds for them as well, but this answer is analytically not satisfactory. In section four, rule individualism is referred to while supplementing the concept of instrumental rationality with the notion of procedural rationality.

In section four I will also come back to the informational content of conventions, as they help agents in forming these expectations in situations of incomplete information. However, the choice of a particular solution to a problem, or, in game theoretic terms, the choice of a particular combination of strategies, still remains to be discussed, and will be done in section three.

---

[24] As Rowe does not differentiate between conventions and (social) institutions, this quotation also holds for what we call conventions.

2.5 Are conventions rational?

In this intermezzo the question as to whether or not rational individuals will conform to a convention is discussed. In the above text, for purposes of clarity the affirmative answer was given, but at the cost of oversimplification. Therefore, a more thorough discussion of the literature on this matter is warranted, which may form the starting point of further research.

This discussion focuses on repeated (pure) coordination problems, because these are the most likely candidates for being solved by a convention. In a repeated coordination problem, the formerly chosen coordination strategy can act as precedent for the present solution. In a new coordination problem this is more difficult, so that the argument runs as follows: if following a convention is not rational in a repeated coordination problem, it will be even less so in a new coordination problem.

For a repeated coordination problem, Lewis (1969, pp.35-36) is one of the first to argue that the formerly chosen solution may be the salient one in a next instance of the coordination problem. He derives his argument from the experiments done by Schelling (to be discussed in section 3.2), in which the configuration of a coordination problem turns out to be the cause of saliency. Lewis is not entirely clear on how choosing the salient solution relates to rationality. This aspect is scrutinized by Gilbert, who points to an assumption, often implicitly made by authors who followed Lewis, that she does not consider valid, namely the assumption that conforming to precedent is inherently rational (Gilbert, 1990, p.8). As Gilbert puts it:

> "... common knowledge of precedent as such will not by itself automatically generate expectation of conformity or conformity on the part of rational agents (Gilbert, 1990, p.10)."
>
> "... the claim that the precedented case will be salient appears to depend on psychological assumptions that may not be true of rational agents as such (ibid., p.11)."

Rationality only implies following a convention if an agent knows that the other player will do likewise. As long as this knowledge is lacking, which is the essence of a coordination problem, rationality does not rule out other behaviour

than the conventional one.[25]

This leads to the conclusion that conformity to a convention does not follow from rationality alone. If agents still succeed in coordinating their actions by way of a convention, "their success depends on a residual element of non-rational action" (Sugden, 1991, p.776). It should be noted that the concept of rationality refers to individual rationality. Gilbert tries to remain within the framework of individual rationality by suggesting a 'Personal Principle Model'. In this model, a player can decide to adopt the strategy rule of sticking to a convention strategy as long as it solves the coordination problem successfully. It is then argued that maintaining the principle can follow from rationality, but the principle itself is arbitrary (Gilbert, 1990, p.13). In my view, this model does not bring theory any further, as it leaves the choice of the initial principle unexplained.

Leaving the concept of individual rationality by substituting other concepts for it, is the essence of two approaches taken in order to underpin the existence of conventions. The first one is taken by Sugden. He suggests to start from the notion of group rationality (1991, p.777). The group of players in a coordination game can be seen as a team that decides on a particular equilibrium strategy. If it does so, all players in the team are better off by following the team strategy than by deviating from it, as long as they know that the other players will conform as well. Players who act on the team rule are then called cooperators (ibid., p.778). In my view, this indicates the essence of the approach: a coordination game is replaced by a cooperative game (see figure 1). The convention is replaced by an institution like: "follow the team rule or thy shalt be thrown out". It may be clear that this approach does not contribute to an explanation of a convention as repeated voluntary coordinating behaviour by agents in a pure coordination problem. The second approach is suggested by Goyal and Janssen (1995), who argue in favor of institutional individualism, which assumes some coordination in order to explain the existence of some other

[25] In the process of rewriting this paper, someone commented "Rules are there, perhaps, because people aren't that rational to conform just because of pay-offs", which is another point to be taken seriously.

coordination. This assumption alleviates the infinite regress problem one faces while starting from scratch, because it allows for some conventions to exist in order to explain other conventions (see also Hodgson, 1993, p.9). Coordination behaviour can be explained by substituting the concept if institutional individualism for a purely individualistic approach. Although the approach may be more intricate than this sketch of it suggests, two questions remain unanswered. The first one is the determination of the intial conventions, which remains arbitrary as this choice cannot be based on fundamentals. The second question is on the relationship between institutional individualism and methodological individualism. At first sight, it seems that adopting the first goes at the expense of the latter. In the remaining text, I will follow the methodological individualism approach as long as possible and see how far one can get with it. This approach may contribute to an answer to the second question.

3. Extensions to coordination games

3.1 Limitations of the game-theoretic framework

The pros and cons of using game theory for the underpinning of conventions and institutions in society are closely linked to the individualistic character of the approach. Explaining group behaviour by analyzing individual decision processes is called methodological individualism. This approach has the advantage of separating norms from actors, that is, norms are not considered as entities (Janssen, 1993, p.xiii). They are the result from agents' interactions, but do not themselves interact with agents. Further, the reduction of the decision process to the individual level gives a clear insight in the extent to which conventions can be explained from this level and the stage at which a model of social interaction is needed (Ullmann-Margalit, 1977, pp.14-17). As games are usually set in a state-of-nature economy, being an economy without conventions or institutions, games may provide a way of endogenizing the emergence of these norms from individually rational behaviour (Schotter, 1981).

This argument can also be put against the game-theoretic framework, because economies are seldom in such a state-of-nature. This limits the usefulness of game theory for analyzing real-world problems. Conventions and institutions such as money, markets, and property rights have a decisive influence on players' behaviour, which cannot be captured by a game, unless they are, rather rigidly, put in as rules of the game, or game theory is extended to incorporate behaviourial assumptions, as is tried by Rabin (in: Friedman, 1994, p.70) for instance. As long as these behaviourial assumptions boil down to a form of (pre-play) communication and result in binding agreements, they are more suitable for cooperative plays than for coordination games, so that conventions still cannot be derived from the game itself (see figure 2 in section 2.1 above). The game format can be used to compare an economy in a 'natural state' with an economy in which some specified conventions or institutions exist, but the development towards these norms cannot be modelled in one game, as then the pay-off pattern changes. That is, strategies that yield highest pay-off in an economy without norms may not do so in an economy with conventions and institutions (Schotter, 1981, p.150-153). This limitation of the game theoretic approach is extremely restrictive for the analysis of conventions, because they often evolve from the specific contextual characteristics of a particular coordination problem: the focal point that will be introduced below.

In the emergence of norms strategic interaction and reciprocal expectations play a crucial role. Strategic interaction exists if the actions agents choose depend on the actions taken by other agents (see section 2.2). Both features can be analyzed within the game format, but the approach is insufficient for pure coordination games. The process of choosing one out of several coordination equilibria that result from the pure coordination game should be modelled as exogenous to the game itself. An individual agent alone cannot decide on a convention. The interaction of agents who form expectations about each other makes simple aggregation impossible. Conventions as the solutions to coordination problems need both a macro and a micro underpinning: the social context influences the individual decision and is influenced by it. The choice of one equilibrium out of several equilibrium outcomes of a game becomes

exogenous to the game instead of endogenous. On the other hand, preferences are now considered endogenous instead of exogenous. Institutions give feedback to agents and influence agents' preferences, because they change the pay-offs that result from strategy combinations. However, if players' preferences change, they become different players. Therefore, the game theoretic assumption that preferences are given, limits the usefulness of game theory for an analysis of institutions.[26] Besides this criticism, Hodgson (1988) also criticizes two other aspects of game theory, namely the assumption that pay-offs are known in advance, and the reduction of complex situations to simple choices of action in games.

3.2 External mechanisms for equilibrium selection

From the above text it can be concluded that game theory does not always give a clue as to which equilibrium to select in a coordination problem with multiple coordination equilibria or a coordination problem with feasible strategy combinations that are inferior in terms of pay-off. Therefore, the need for an additional type of decision criterion arises: "coordination requires the common acceptance of some sort of suggestion" (Schelling, 1960, p.144). This suggestion should indicate which strategy is to be chosen. In 1960 already, Schelling introduced the idea of focal points as a source of suggestion:

> "People can often concert their intentions or expectations with others if each knows the other is trying to do the same. Most situations (...) provide some clue for coordinating behavior, some focal point for each person's expectation of what the other expects him to expect to be expected to do." (Schelling, 1960, p.57)

This source of suggestion can be fairly arbitrary as long as it satisfies the requirement of being shared among the participants of the coordination problem. Three sources of suggestion are relatively often shared by the participants, those being contextual features, rules of thumb and past period outcomes. I will discuss these in turn.

---

[26] For more elaboration on this point, see Hodgson (1988, pp.134-7; 1993, pp.7-9).

The ‘clue for coordinating behaviour’ may be given by the configuration of the problem (ibid., p.69) without any reference to external factors. An often-cited example is the determination of a meeting point by people who cannot communicate about a place to meet. Uniqueness or obviousness are crucial characteristics of a solution. As these cannot be decided upon by using objective criteria, intersubjectivity is decisive. This implies that the focal point is endogenous, because it is determined by the problem configuration. In this respect, this first source of suggestion differs from the other two, as rules of thumb are determined outside the problem and history is separated from the problem in the time dimension. Rules of thumb serve as a focal point if they are shared by a group of agents. They are useful in situations with too much or too little information to go through an optimization procedure, in situations of time constraints and in situations in which only little or infrequent communication is possible. One can think of oligopolists who use a common mark-up pricing rule, a rule to determine their advertisement budgets as a percentage of gross revenue, or the same indicators to determine when to change prices. Shared rules of thumb give information on the predictability of competitor’s behaviour and may thus contribute to orderly markets. Focal points can also be given by common habits, such as issuing new prices once a year, and retailer’s discounts in January.[27] Habits emerge if for several periods of time the past choices determine the actual ones, so that history determines the choices made today.[28] This leads to path dependency, which will be discussed more extensively in the next section. Another way in which past period outcomes determine the

[27] A nice example is found in the American automobile industry. In the early nineteen-thirties the seasonal fluctuations in employment used to be large because of the introduction of new models and the spring burst of demand associated with good weather being at the same time. Changing the timing of the yearly automobile show provided a solution, as production of the new models could now take place earlier in the season (Cooper and Haltiwanger, 1993).

[28] One may object against this, as habits do not seem compatible with rationality. In terms of adaptive behaviour, however, past period choices may (partly) determine the actual ones. This underpins section 2.5’s conclusion that the framework of rationality may be too restrictive.

equilibrium selection in a coordination problem, is by norm messages (Schotter, 1981, pp.128, 134-9). Institutions are then created as a result of updates of information sets, so that players know what choices have been made on previous occasions. Present players then need only know what choice is thus prescribed by the institution. This situation has the advantage that only a limited amount of information is needed for deciding on an equilibrium (ibid., p.139), but it can easily lead to inflexibility and hysteresis. This ends the discussion of the three sources of suggestion.[29]

Perhaps, conventions and institutions fit in the context of an Overlapping Generations model, if these norms are interpreted as an asset that is given over from each generation to the next.[30] The reason they can be seen as an asset to an economy is that they save the costs of information collection and processing and the effort of collectively making a decision as they prescribe equilibrium selection behaviour in coordination problems. To make this more concrete, imagine a market place at which stands can be hired for the yearly fair. Every year there will be new marketeers and more experienced ones. The latter know of conventions, for instance on when to start discounting. They can tell this to their new colleagues, so that a yearly discussion in order to collectively decide on this is prevented. The new marketeers will outlive the elderly ones and instruct the new younger generation. Even if at some time the original problem behind the convention is not known any longer, the convention remains to serve as a solution to it (Schotter, 1981, p.14).

If the source of suggestion creates an asymmetry between players, a convention can develop even easier, as is described by Sugden (1986). The development of a convention as described by Sugden can be summarized as follows. A game is played by two players. If a player considers his opponent as different from himself and if he considers this difference as crucial to the pay-

[29] Bacharach (1993) and Janssen (1995) discuss the rationalizability of focal points.

[30] However, this approach may bring forth the same indeterminacies as do other Overlapping Generations Models. Whether or not this can be considered an advantage in terms of explanatory value remains to be studied.

offs to both players, the game is called asymmetric. Therefore, in an asymmetric game a player will first determine which of the two possible roles he is to play, before deciding on his behaviour. If both players recognize the asymmetry and if they behave accordingly, a convention can develop. The crux of this evolutionary process is the mutual recognition of the characteristics of the players that determine the asymmetry. In other words, the players somehow must manage to agree beforehand on who is to play what role. The characteristics that are crucial are often not rationally determined, but seem to be a matter of prominence: features that seem unique, prevalent, or easy to recognize in other situations. The easier a player can conjecture which aspect the other player considers decisive on the division of roles, the easier a convention emerges on how to play a particular game. It is the recognition of asymmetry that prevents players from a situation of infinite regress, because the strategy bringing the highest pay-off to the players is now contingent on their role instead of the action of the other player. Communication supports and accelerates the evolution of conventions, because information diffuses more quickly among players, but it is not necessary for the process as such. During the process of broader acceptance of a convention in society, the asymmetry becomes larger, thereby accelerating the acceptance: the development of conventions is a self-reinforcing process. Sometimes more conventions compete for a period of time, until the one that is most widely used, becomes the standard. For the evolution of conventions to be possible, one must assume that games are played repeatedly by anonymous[31] players who are able to learn from past behaviour.

Before some notions from the New Institutional Economics follow in the next section, a preliminary assessment is given. Three types of coordination games are distinguished, which differ in the degree of coordination versus conflict. Game theory turns out to be insufficient for solving pure coordination problems. The outcomes call for a norm that is external to the game. In case of a pure coordination game, this norm will be a convention and secures the

[31] Anonymity is a condition that prohibits reputation building and forces the players to act according to their role instead of their knowledge of their opponent.

attainment of a coordination equilibrium. A PD-game can be extended with an institution, which forces the players of it to choose the strategy combination which delivers the highest total pay-off. For bargaining games, the norm may be of a convention type or an institution type, depending on the initial strategy combination that is chosen by the players. The way coordination norms come into existence and are disseminated in an economy is explained within the framework of instrumentally rational agents.

In the next section the game theoretic interpretation of coordination problems is complemented with a New Institutional Economics point of view, which focuses on information and transaction costs. This approach differs from the game-theoretic state-of-nature approach in that it first takes institutions, which is also to include conventions, as given and then tries to rationalize their existence.

4. Information and institutions: some notions from New Institutional Economics

New Institutional Economics explicitly aims at explaining the existence of institutions. As this paper surveys theories on conventions and institutions that facilitate coordination, it is incomplete without discussing this branch of economic theory. Further, the notions from New Institutional Economics are complementary to the ideas described above. Before this is shown, the adjective ’New’ needs to be motivated. New Institutional Economics differs from ‘Old’ Institutional Economics in several respects: it takes individuals as exogenous, it satisfies methodological individualism,[32] it tries to underpin the emergence of institutions and their comparative efficiency and it develops new concepts of markets and firms. As such, the New Institutionalists’ approach closely resembles the treatment of coordination problems above. What is a possibly more attractive feature of ‘Old’ Institutional Economics, though equally more

[32] On methodological individualism, more is said in sections 2.5. and 3.1.

complex, is that it explicitly takes into account the feedback relation of the institutional context on an individual agent and can be more easily combined to evolutionary economics (Hodgson, 1993). By not satisfying the criterion of methodological individualism, Old Institutional Economics may be closer to the institutional individualism approach that has been mentioned in section 2.5.

In the New Institutional approach the term institutions has a different connotation as compared to the above text. In the former sections, institutions were norms with an external sanction mechanism that made the norm a binding one. Besides, conventions were norms without the need for such a mechanism, as they were self-enforcing. The New Institutionalists choose a broader interpretation that covers both conventions and institutions in the sense of coordination norms, so these terms are now treated as synonymous. A definition by North follows soon below.

From the Coase theorem it can be inferred that institutions do not matter as long as transaction costs are zero, because the institutional setting can then be changed at zero costs.[33] According to North, two sorts of transaction costs exist, namely measurement cost and enforcement cost. This implies that institutions matter: the initial institutional setting, at least partially, determines the outcomes of behaviour. Institutions can then be defined as "the humanly devised constraints that structure human interaction" (North, 1994, p.360). They consist of formal rules, informal norms, and enforcement characteristics which together act as constraints on the human behaviour. It may now be clear that both types of coordination norms that were used distinctly in the above text, satisfy this definition of institutions, so that the discussion may safely be continued by only referring to institutions in the sense of North' definition. This approach is a methodological individualistic one as far as it focuses on the individual instrumentally rational agent who maximizes his utility while being constrained by his context in which institutions may exist. However, as North admits, not all existing institutions can be explained by the transaction costs

[33] Some authors see property rights as an institution. For the Coase theorem to hold, property rights must be well-defined.

approach (North, 1990, p.7). In other words, some institutions are not the most efficient solution to the problem they were intended to solve or the need they were intended to satisfy. This point has also been made in the section on coordination games, where conformity to a coordination norm is considered rational even if the norm itself is not necessarily so.

This limitation of the instrumental rationality approach calls for an extension of theory, which can go (at least) two ways. The first approach is an evolutionary one: individuals need not be instrumentally rational, but the instrumentally rational behaviour is the attitude that is fittest in the evolutionary struggle for survival, in this case being economic competition (North, 1990, p.19). However, even if this is so, the existence of less efficient institutions in the long run cannot be explained. The second approach aims at a different concept of rationality. "Procedural rationality is the way that the institutional and informational contribution, which cannot itself be reduced to instrumental rationality, is introduced into the explanation at the level of the individual" (Hargreaves Heap, 1989, p.71). This concept "portrays the individual as a rule follower, a person who follows norms, recipes or procedures for action" (ibid., p.4). It closely resembles Rowe's concept of rule individualism that was introduced in section 2.4. Rule individualism and procedural rationality both consider the rationality of a pattern of action instead of the actions themselves. This can be considered an extension of rationality as opposed to the concept of bounded rationality, which is a limited sort of rationality (Hargreaves Heap, 1989, p.119). The result will, if not maximal, be satisfactory at least, so that the outcome with an institutionalized rule of behaviour is better than without one (Schotter, 1981, p.148-9). Procedural rationality contributes to a macrofoundation of microeconomics to the extent that procedures are common among agents. In this view, procedural rationality is not an extension of instrumental rationality, but a reversal of it, as instrumental rationality underpins a microfoundation for macroeconomics (Hargreaves Heap, 1989, p.146).

Procedural rationality mitigates lack of information, because it makes actions by individual agents more predictable. This effect is even stronger if procedures are institutionalized, so that the mere existence of the institution aids

in forming expectations on agents' actions. In this sense, shared mental models can be considered institutions (Denzau and North, 1994). The cost of this predictability is limited individual freedom, which is in line with the North definition of institutions as humanly devised constraints on human behaviour.[34] All this only holds in a society with incomplete information, since otherwise, by definition, no need for extra information can exist: in a world with perfect information, conventions are redundant. Incomplete information forces agents to act on predictions in case information is lacking, and extra information that makes predictions better enhances efficiency. As prices are information carriers (cf. Alchian, 1977), institutions can be seen as complementary to prices: "Social and economic institutions are informational devices that supplement the informational content of economic systems when competitive prices do not carry sufficient information to totally decentralize and coordinate economic activities" (Schotter, 1981, p.109). This complementarity can be put on a scale, with perfect competition at one side and monopoly on the other. In a perfectly competitive economy, prices contain all relevant information; in an oligopoly situation prices are complemented by conventions; and in a monopoly conventions supply information as prices do not reflect production cost information (Schotter, 1981, p.158).

In the above discussion of institutions and information, the point was made that institutions enhance the predictability of action and thereby contribute to information in society. More can be said about this relationship if the influence of incomplete information or fundamental uncertainty on the existence of institutions is considered. To give an impression, two possible underpinnings of this relationship are introduced. Both are based on an amendment to the (instrumentally or procedurally) rational agent insofar as now allowance is made for limitations on the cognitive capacity of agents, which brings me closer to reality. First, in situations of fundamental uncertainty agents are not able to optimize over the whole range of possible actions. Their individual limited

[34] The reader may note that what is called 'the institutionalization of procedures' in this section, is essentially the same as 'repeated patterns of behaviour becoming a binding norm with a sanction mechanism' in former sections.

cognitive capacity forces them to restrict the amount of information they absorb and to simplify their behaviour. Predictable behaviour then follows from uncertainty that prevents agents to maximize (Heiner, 1983, p.561). This predictable behaviour may influence the expectations agents form about each other's behaviour and expectations, so that it may be worthwile to continue behaving in a predictable way, which makes the behaviour binding so that it satisfies the definition North gives of institutions. Second, in situations of uncertainty, beliefs and trust become at least as important as knowledge. Institutions then can become carriers of these beliefs and are able to sustain because they are believed to do so. In section 5, central banks and the backing of fiat money by gold are given as examples of such institutions (Van Ees and Garretsen, 1994, p.20).

Another relationship between institutions and information runs as follows. In a context of incomplete information, transaction costs in the form of information costs exist. These transaction costs reflect market imperfections, and result in increasing returns to scale for institutions, so that larger institutions require less information per 'participating agent' than do smaller institutions. This may result in the existence of only a few large institutions that disperse more information in terms of predictable behaviour. Institutions that are established tend to persist because of this scale effect. In terms of procedural rationality this argument can be put as follows: if many incompletely informed agents conform their behaviour to a common procedure, it is rational to join them (cf. North 1990, p.108). This argument resembles the line of thought expressed in the treatment of pure coordination problems above.

Here, the notion of path dependency comes in: the existence of institutions not only depends on their efficiency or utility increasing effects, but on their existence in former times as well.[35] If initial conditions are absent, this leads to indeterminacy. With the presence of initial conditions, an infinite regress problem may emerge, because at every stage in the development of an

[35] In order to capture this idea, time should be seen as historical instead of logical: events are dependent on the moment in which they happen (see O'Driscoll and Rizzo, 1985).

institution, the preceding stage is a necessary element of an underpinning of its existence. Perhaps, for the theorist only one approach remains, which is taking some initial conditions as given, however dissatisfactory this approach may be. I will shortly come back on this in the concluding section.

In an economy with incomplete information, learning may improve the agents' ability to coordinate. The concepts of historical time and path dependency are linked with learning, because learning is irreversible (one cannot wipe out human routines and memory) and because the knowledge that results from past learning determines the present learning process. The latter statement refers to framing and comes from cognitive science (see e.g. Denzau and North, 1994). Here, learning is described as the interpretation of information within an agent's framework, which serves as his model of the world, as well as redesigning the framework itself if new information requires so (representational redescription). A parallel can be drawn between the evolution processes of mental frameworks at one side, and conventions and institutions at the other, insofar as all three tend to persist once they emerge, but they are replaced by new ones once their context changes drastically.[36] They are, so to speak, locally robust. A drastic change of context refers to, respectively, different (sorts of) information, and new problems, new types of behaviour or a new focal point. The idea of norms being locally robust resembles the 'corridor' designed by Leijonhufvud. Within this 'corridor' around a full-coordination time-path, deviations do not challenge the stability of the path, whereas disturbances that bring the economy outside the 'corridor' lead to a fundamental different state of the economy, with Keynesian characteristics (Leijonhufvud, 1981, pp.109-110). In the same fashion a coordination norm may change once the coordination problem or its context has changed significantly. Learning can also be a source of change for institutions if agents learn of a substantially more efficient institution to perform a task. The same holds for a convention shift: the replacement of one convention by another. Shared mental models form a culture

[36] This notion is comparable to Leibenstein's inert areas: "a set of bounds, say upper and lower bounds, within which any decision maker will not change his decision" (Leibenstein, 1984, p.79).

or an ideology and this may contribute to the explanation of different institutional configurations in different countries, for instance Islamic banking without interest versus Western banking with interest.[37]

New Institutional Economics provides many building blocks for an interpretation of institutions. It brings in the necessary extra's that game theory misses because of its rigorous assumptions, especially the assumption of a state-of-nature economy. At the other hand, this very assumption makes game theory attractive, because it can be used without having to make arbitrary assumptions about the initial institutional setting of an economy with coordination problems. In this way, both theories are incomplete and complement each other. In the next section, they will be given a try in a preliminary investigation of their applicability to examples of real-life coordination problems.

5. Application: examples of coordination problems

In the introduction of this article coordination problems were said to occur in situations of incomplete information. A game theoretic discussion followed, in which even completely informed agents were unable to solve coordination problems. Coordination norms were introduced as information carriers that were exogenous to the coordination problem itself. This last section returns to the context of incomplete information that was started with in the introduction. By way of examples of coordination problems, an exploration is made of the possibilities of using the theoretical notions introduced above. These concepts may thus contribute to an underpinning the existence of conventions, institutions and bargaining norms in an economy. At this stage, only indications are given. A more detailed study may start from here.

Agents who have only limited information available will not always be able to reach maximal utility, even if they are rational in the instrumental sense.

[37] For other illustrations and more on cultural institutional diversity see North (1991) and Denzau and North (1994).

Procedural rationality provides a way of coping with this, because rules for behaviour can support a desirable coordination outcome, as was described above. Conventions and institutions provide these rules of behaviour. Such coordination norms are seen in the labor market, where employers and workers compare wages across firms in their industry and use this information as a reference wage. Then, the reference wage becomes a focal point. If both employers and workers consider it reasonable to base their wage offer and wage demand, respectively, on such a reference wage, this is a kind of implicit contract which functions as a convention. This behaviour may easily lead to inertia, and may even more do so if the reference wage is based on information from past periods (Okun, 1981, p.95).

If agents have to do without complete information to base their decisions upon, their trust becomes an alternative criterion. For instance, people start holding buffer stocks of money if they lack information on transaction opportunities. They will also do so if information is costly to attain and a trade-off between buying information and possibly missing profitable transaction opportunities exists (Laidler, 1990, p.12). By keeping buffer stocks of money, people put their trust in the monetary system and in the central bank that runs it (Van Ees and Garretsen, 1994, p.20). As a monetary system only works if all, or at least a critical part of, the agents use it, a coordination problem is present. To the individual agent, the system is only trustworthy if many other agents put enough trust in it to use it. The coordinating norm is partly a convention, because using money lowers transaction costs and is therefore self-enforcing, and partly an institution as issuing money is reserved to the central bank.

In situations of incomplete information, agents can learn.[38] From an economic point of view, learning can be seen as the acquisition and processing of information. An instrumentally rational agent will learn until the marginal costs of learning equal the marginal benefits gained from it. In situations of incomplete information however, one may not possess enough information in

[38] In section 4 the process of learning has been described, as well as its connection with historical time.

order to determine the costs and benefits of further learning, so the agent must estimate them. This is seen in customer markets, where agents shop around for favourable offers by suppliers until they conjecture having enough information to decide on a purchase (Okun, 1981). Their shopping behaviour aims to minimize the risk of missing a relatively favourable offer, with time spent on shopping as a constraint. The supplying firm, at the other hand, has a stake in a long-term relationship with its customers and will put effort in setting up such a relationship (Clower and Howitt, 1994, p.9). To the extent that coordination norms provide information, as was discussed in section 4, one sees them in customer markets as well. Perhaps reputation building serves as an example, for it contains an implicit promise by a supplier to behave in conformity to the reputation, e.g. charging low prices, and an implicit promise by customers to buy at that particular shop.

Perhaps coining can be interpreted as an institution that removes uncertainty with respect to quality and weight of the precious metal used as a means of exchange. Coins[39] that are declared legal means of payment are institutions because falsifying these coins is prohibited and sanctioned. At the same time, the choice of any commonly accepted means of exchange is a pure coordination problem because of the network externalities involved, so that the use of a particular type of coins is a convention. It may even be possible to interpret Gresham's Law from the point of view of competing conventions.

Another occasion of a coordination problem occurs with the selection of a currency in a process of monetary integration of countries. This bargaining problem results from the common interests of using the same currency and so enjoying a reduction in transaction costs at the one hand, and the conflict of not wanting to be the only country that has switched to using the new currency. As an example, the introduction of a European common currency, called ECU, as a bargaining game with Germany and The Netherlands as players is given:

[39] Most of this argument holds for banknotes as well.

(table 4)

		Germany	
		DMark	ECU
Netherlands	Dfl	0 , 0	0 , -1
	ECU	-1 , 0	1 , 1

In this game it can be seen that switching both (ECU, ECU) is most profitable for both players as compared to the initial state (Dfl, DMark) with two currencies. The country that switches alone, is worst off, because no gains are set against the switching costs that are made by it. In order to ensure both countries that their counterpart is committed to switching, a binding agreement is necessary: cooperation takes place in the institute of the European Central Bank.

Bargaining norms, being the hybrid in-between type of coordination norms, play a role at the timing of sales, the type of after-sales service, reasons for price increases that are considered legitimate, the timing of issuing new price lists, the closing times of shops, etcetera (see e.g. Kahneman, Knetsch and Thaler (1986), Blinder (1991) and Leibenstein (1984)). These examples contain a conflict element to the extent that they are instruments for competition. For example, the oligopolist who first starts with sales may attract most customers. They are also subject to coordination, because all oligopolists are best off with synchronized behaviour. As long as there is no ‘objective’ criterion for deciding upon these matters, agreement must be reached on a focal point. This agreement cannot be reached by communication, because anti-cartel legislation forbids so, so the focal point must emerge from the context in which the oligopolists operate (Hargreaves Heap, 1992, p.122). Then, the focal point needs to be ’obvious’, prominent or otherwise unique and serves as an anchor for mutual expectations of the oligopolists. If the coordination norm is established, it provides information on how to behave in the oligopolistic market.

Comparing the use of money and price setting in oligopolistic markets, the use of money may be considered the outcome of a cooperative game, because democratically created legislation enforces taxes to be paid with it, and price

setting behaviour as a strategy in a non-cooperative, or coordination game,[40] because a binding agreement is prohibited by anti-cartel legislation. However, the difference is not that sharp. Falsifying the legal money may be forbidden, the use of legal money is not prescribed. Therefore, the possibility of barter or creating 'local money'[41] offers an escape opportunity and makes the use of money less of a cooperative type of game since, in that case, the coordination element is most important. On the other hand, price setting agreements in an oligopoly, however tacitly made, may be binding in situations of credible threats of price wars. Then, a binding agreement exists, and the price setting game is cooperative.

[40] For their equivalence, see section 2.1.

[41] At present, successful experiments are done with LETS (Local Exchange Trading Systems) in Australia and New Zealand. In Great Britain and the Netherlands, similar systems have been started up (Aktie Strohalm, 1994).

References

Aktie Strohalm, 1994, *Denken en doen durven omkeren*, Utrecht

Alchian, A.A., 1977, Why Money?, *Journal of Money, Credit and Banking*, vol.9, pp.133-140

Bacharach, M., 1993, Variable Universe Games, Chapter 13 in: K. Binmore, A. Krinan and P. Tani (eds.), *Frontiers of Game Theory*, Cambridge Mass.: MIT Press

Blinder, A.S., 1991, Why are Prices Sticky? Preliminary Results from an Interview Study, *NBER Working Paper No.3646*, Cambridge

Boyer, R. and A. Orléan, 1992, How do Conventions Evolve?, *Journal of Evolutionary Economics*, vol.1(2), pp.165-177

Cooper, R. and A. John, 1988, Coordinating Coordination Failures in Keynesian Models, *Quarterly Journal of Economics*, Vol.CIII(3), pp.441-463

Cooper, R. and J. Haltiwanger, 1993, Automobiles and the National Industrial Recovery Act: Evidence on Industry Complementarities, *Quarterly Journal of Economics*, Vol.108, pp.1043-1071

Clower, R.W. and P. Howitt, 1994, Taking Markets Seriously: Groundwork for a Macrofoundation, *paper for the Eastern Economics Association Meetings*, Boston

Denzau, A.T. and D.C. North, 1994, Shared Mental Models: Ideologies and Institutions, *Kyklos*, Vol.47(1), pp.3-31

Friedman, J.W., (ed.), 1994, *Problems of Coordination in Economic Activity*, Boston

Furth, D., 1989, A General Theory of Equilibrium Selection in Games: A Review Article, *Journal of Institutional and Theoretical Economics*, Vol.145, pp.730-738

Furth, D., 1993, Game Equilibrium Modelling, *De Economist*, Vol.141(3), pp.353-379

Gilbert, M., 1990, Rationality, Coordination and Convention, *Synthese*, Vol.84, pp.1-21

Goyal S. and M.C.W. Janssen, 1995, Can we Rationally Learn to Coordinate?, *Theory and Decision*, forthcoming

Hargreaves Heap, S.P., 1989, *Rationality in Economics*, Oxford

Hargreaves Heap, S.P., 1992, *The New Keynesian Macroeconomics*, Aldershot

Hargreaves Heap, S.P., 1994, Institutions and (Short-Run) Macroeconomic Performance, *Journal of Economic Surveys*, Vol.8(1), pp.35-56

Harsanyi, J.C. and R. Selten, 1988, *A General Theory of Equilibrium Selection in Games*, Cambridge

Heiner, R.A., 1983, The Origin of Predictable Behaviour, *American Economic Review*, vol.73(4), pp.560-595

Hodgson, G.M., 1988, *Economics and Institutions: a Manifesto for a Modern Institutional Economics*, Cambridge

Hodgson, G.M., 1993, Institutional Economis: Surveying the 'Old' and the 'New', *Metroeconomica*, Vol.44(1), pp.1-28

Janssen, M.C.W., 1993, *Microfoundations: A Critical Inquiry*, London

Janssen, M.C.W., 1995, *Rationalizing Focal Points*, Tinbergen Institute Research paper No.TI95-35, Rotterdam

Kahneman, D., J.L. Knetsch and R. Thaler, 1986, Fairness as a Constraint on Profit Seeking: Entitlements in the Market, *American Economic Review*, Vol.76(4), pp.728-741

Keynes, J.M., 1936, *The General Theory of Employment, Interest and Money*, reprinted in 1957, London: MacMillan

Kreps, D.M., P. Milgrom, J. Roberts and R. Wilson, 1982, Rational Cooperation in the Finitely repeated Prisoners' Dilemma, *Journal of Economic Theory*, Vol.27, pp.245-252

Kreps, D.M., 1990, *A Course in Microeconomic Theory*, New York

Kuipers, S.K., H. van Ees, H. Garretsen, G.H. Kuper, E. Sterken, and E. Wester, 1994, *Evenwicht, Conjunctuur en Groei; Moderne macro-economische theorievorming*, Utrecht

Laidler, D., 1990, *Taking Money Seriously and Other Essays*, New York

Leibenstein, H., 1984, On the Economics of Conventions and Institutions: an Exploratory Essay, *Journal of Institutional and Theoretical Economics*,

Vol.140(1), pp.74-86
Leijonhufvud, A., 1981, *Information and Coordination*, Oxford
Lewis, D.K., 1969, *Convention, A Philosophical Study*, Cambridge
North, D.C., 1990, *Institutions, Institutional Change and Economic Performance*, Cambridge
North, D.C., 1994, Economic Performance Through Time, *American Economic Review*, vol.84(3), pp.359-368
O'Driscoll, G.P. and M.J. Rizzo, 1985, *The Economics of Time and Ignorance*, Oxford
Okun, A.M., 1981, *Prices and Quantities: a Macroeconomic Analysis*, Oxford
Rasmusen, E., 1989, *Games and Information, An Introduction to Game Theory*, Oxford
Rowe, N., 1989, *Rules and Institutions*, New York
Schelling, T.C., 1960, *The Strategy of Conflict*, Cambridge
Schotter, A. 1981, *The Economic Theory of Social Institutions*, Cambridge
Sugden, R., 1986, *The Economics of Rights, Co-operation and Welfare*, Oxford
Sugden, R., 1989, Spontaneous Order, *Journal of Economic Perspectives*, Vol.3(4), Fall 1989, pp.85-97
Sugden, R., 1991, Rational Choice: A Survey of Contributions from Economics and Philosophy, *The Economic Journal*, Vol.101, pp.751-785
Ullmann-Margalit, E., 1977, *The Emergence of Norms*, Oxford
Van Damme, E.E.C. and A. Heertje, 1994, Speltheorie in beweging, *Economisch Statistische Berichten*, 19-10-1994, pp.936-941
Van Ees, H. and H. Garretsen, 1994, *Existence and Stability of Conventions and Institutions in a Monetary Economy*, mimeo, Groningen
Witteveen, H.J., 1994, *Toespraak voor de Koninklijke Vereniging voor de Staathuishoudkunde op 10 december 1994*, Amsterdam
Wubben, E.F.M., 1993, *Markets, Uncertainty and Decision-making*, Rotterdam